

Multi-copy programmable discrimination of general qubit states

G. Sentís, E. Bagan, J. Calsamiglia and R. Muñoz-Tapia

*Física Teòrica: Informació i Fenòmens Quàntics, Facultat de Ciències, Edifici Cn,
Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona) Spain*

Quantum state discrimination is a fundamental primitive in quantum statistics where one has to correctly identify the state of a system that is in one of two possible known states. A programmable discrimination machine performs this task when the pair of possible states is not a priori known, but instead the two possible states are provided through two respective program ports. We study optimal programmable discrimination machines for general qubit states when several copies of states are available in the data or program ports. Two scenarios are considered: one in which the purity of the possible states is a priori known, and the fully universal one where the machine operates over generic mixed states of unknown purity. We find analytical results for both, the unambiguous and minimum error, discrimination strategies. This allows us to calculate the asymptotic performance of programmable discrimination machines when a large number of copies is provided, and to recover the standard state discrimination and state comparison values as different limiting cases.

PACS numbers: 03.67.Hk, 03.65.Ta

I. INTRODUCTION

Discrimination between given hypotheses is one of the most basic tasks in our every day lives. Very often we are confronted with the necessity of having to identify an option between some possible choices based on some acquired evidence. In the quantum setting the discrimination problem consists of identifying one of two possible states given a number of identical copies available for measurement. This task encompasses a plethora of non-trivial theoretical and experimental implications. In the usual setting the a priori states are known, i.e., the classical information characterizing the possible states is provided and the discrimination protocol is tailored for this specific information. One usually considers two types of approaches: unambiguous [1] and minimum error [2] discrimination. An unambiguous protocol is one where the identification of the state is error free. Of course, this is only possible stochastically, i.e., unless the states are orthogonal, the protocol must give an inconclusive answer (the “I do not know” outcome) with a non vanishing probability. In the minimum error approach, the protocol always yields a definite answer, which may be wrong some of the times. An optimal protocol is one which minimizes the inconclusive or the error probability. It may also be possible to go continuously from one case to the other by considering margins of error probabilities [3]. In spite of being such a fundamental problem, only very recently a closed expression for the asymptotic error probability has been obtained (see [4–6] and references therein), the quantum Chernoff bound, from which metric distances and state densities [7] can be derived.

Very much in the spirit of universal computers, it is interesting to consider discrimination devices that are not specialized in a specific discrimination instance but can discriminate between arbitrary pairs of states [8, 9]. In these, the set of possible states enter the device as “programs”, i.e., the classical description of the states is not provided beforehand, rather the information is incorpo-

rated in a quantum way (this can also be viewed as an instance of relative information [10]). These devices have program ports that are loaded with the program states, and a data port that is loaded with the unknown input state one wishes to identify. The device will identify the state of the data port as being one of the states fed in the program ports, but this identification will in general be erroneous with a probability that decreases with the number of copies of the states entering the ports. One can also regard these devices as learning machines [11], where the device is instructed through the program ports about different states, and based on this knowledge the machine associates the state in the data port with one of the states belonging to the training set. Increasing the number of copies of states at the program and data ports of course increases the chances of correct identification.

It is particularly relevant to understand how the probability of error scales with an increasing number of copies and what are the corresponding error rates. The value of this rate is one of the most relevant parameters assessing the performance of the device. We will consider the discrimination of two general qubit states, although most of our results can be generalized to higher dimensional systems (see [12, 13] for a single copy continuous variable setting). For simplicity we will assume that the prior occurrence probability of each state is identical and compute the unambiguous and minimum error rates for optimal programmable devices.

We first study the performance of such devices for pure states. We compute the error probabilities for any number of pure qubit states at the input ports. Some of the results are already available in the literature [9, 14–18], but the way we formalize the problem here is crucial to treat the more general mixed state case. In addition we obtain analytical expressions that enable us to present the results and study limiting cases in a unified way. In particular, when the program ports are loaded with an infinitely large number of copies of the states we recover the usual state discrimination problem [2], since

it is clear that then one has the classical information determining the states entering the program ports. On the other hand, when the number of copies at the data port is infinitely large, while the number of copies at the program ports are kept finite, we recover the state comparison problem [19, 20].

We extend the previous pure state study to the case of mixed input states. In this scenario we only compute the minimum error probability, as no unambiguous answers can be given if the states have the same support [21]. The performance of the device for a given purity of the input states allows to quantify how the discrimination power is degraded in the presence of noise. The expressions here are much more involved, however one can still exploit the permutation symmetry of the input states to write the problem in a block-diagonal form [22, 23]. We then obtain closed expressions for the probability of error that can be computed analytically for small number of copies and numerically evaluated for a fairly large number of copies. We are also able to obtain analytical expressions for some asymptotic rates. Again, the leading term, as in the pure state case, is seen to coincide with the average minimum error for known states.

We also analyze the fully universal discrimination machine, i.e., a device that works optimally for completely unknown input states. In this case one has to assume a uniform distribution for the purity. In contrast to the pure state distribution, there is no unique choice [26], and different reasonable assumptions lead to different uniform priors. Here we consider hard-sphere, Bures and Chernoff priors.

The paper is organized as follows. In the next Section we obtain the error probabilities for pure states when each program port is fed with n copies of each state and there are m copies of the unknown state entering the data port. In Sec III we study the asymptotic rates in several scenarios. In Section IV we analyze the performance of these devices when the ports are loaded with copies of states of known purity and obtain some interesting limiting cases in Sec. V. We finally obtain the error rates for the fully universal programmable machine. Some brief conclusions follow and we end up with two technical appendices.

II. PURE STATES

Let us start by fixing the notation and conventions used throughout this paper. We label the two program ports by A and C . They will be loaded with states $|\psi_1\rangle$ and $|\psi_2\rangle$, respectively. The data port, B , is the middle one and will be loaded with the states we wish to identify. We also use the short hand notation $[\psi]$ to denote $|\psi\rangle\langle\psi|$ and similarly $[\psi\phi\dots] = [\psi] \otimes [\phi] \otimes \dots = |\psi\rangle\langle\psi| \otimes |\phi\rangle\langle\phi| \otimes \dots$. We may also omit the subscripts A, B and C when no confusion arises. We assume that the program ports are fed with n copies of each state and the data port with m copies of the unknown state. This

is a rather general case for which closed expressions of the error probabilities can be given. The case with arbitrary n_A, n_B , and n_C copies at each port is discussed in appendix A. The expressions are more involved but the techniques are a straightforward extension of the ones presented here.

When the state at the data port is $|\psi_1\rangle^{\otimes m}$ or $|\psi_2\rangle^{\otimes m}$, the effective states entering the machine are given by the averages

$$\begin{aligned}\sigma_1 &= \int d\psi_1 d\psi_2 [\psi_1^{\otimes n}]_A [\psi_1^{\otimes m}]_B [\psi_2^{\otimes n}]_C \\ \sigma_2 &= \int d\psi_1 d\psi_2 [\psi_1^{\otimes n}]_A [\psi_2^{\otimes m}]_B [\psi_2^{\otimes n}]_C,\end{aligned}\quad (1)$$

respectively. The integrals can be easily computed using Schur lemma, $\int d\phi [\phi]_X = \frac{1}{d_X} \mathbb{1}_X$, where d_X is the dimension of the Hilbert space spanned by $\{|\phi\rangle\}$ and $\mathbb{1}_X$ is the projector onto this space. Hence

$$\begin{aligned}\sigma_1 &= \frac{1}{d_{AB}d_C} \mathbb{1}_{AB} \otimes \mathbb{1}_C \\ \sigma_2 &= \frac{1}{d_A d_{BC}} \mathbb{1}_A \otimes \mathbb{1}_{BC},\end{aligned}\quad (2)$$

where $\mathbb{1}_{XY}$ is the projector onto the completely symmetric subspace of $\mathcal{H}_X \otimes \mathcal{H}_Y$ and $d_{XY} = \text{tr} \mathbb{1}_{XY}$ is its dimension. For qubits we have $d_A = d_C = n + 1$ and $d_{AB} = d_{BC} = n + m + 1$.

The structure of the states (2) suggests the use of the angular momentum basis, $|j_A, j_B(j_{AB}), j_C; JM\rangle$ for σ_1 and $|j_A, j_B, j_C(j_{BC}); JM\rangle$ for σ_2 . The quantum numbers $j_{AB} = j_A + j_B$ and $j_{BC} = j_B + j_C$ recall the way the three spins are coupled to give the total angular momentum J . Here the angular momenta have a fixed value determined by the number of copies at the ports, $j_A = j_C = n/2$, $j_B = m/2$. So, we can very much ease the notation by only writing explicitly the labels j_{AB} and j_{BC} . We would like to stress however that in general one needs to keep track of all the quantum numbers, specially when dealing with mixed states as in Sec. IV.

In σ_1 the first $n + m$ spins are coupled in a symmetric way, while in σ_2 the symmetrized spins are the last $n + m$, thus $j_{AB} = (n + m)/2 = j_{BC}$. The states are diagonal in the angular momentum bases discussed above, and we have

$$\begin{aligned}\sigma_1 &= \frac{1}{d_{AB}d_C} \sum_{J=0,1/2}^{m/2+n} \sum_{M=-J}^J [j_{AB}; JM] \\ \sigma_2 &= \frac{1}{d_A d_{BC}} \sum_{J=0,1/2}^{m/2+n} \sum_{M=-J}^J [j_{BC}; JM],\end{aligned}\quad (3)$$

where the lower limit of the first summation takes the value 0 (1/2) for m even (odd). Notice that the spectrum of both matrices is identical and that the basis elements of their support differ only in the way the three spins are coupled. Further, the key feature of the total angular momentum bases is the orthogonality relation

$$\langle j_{AB}; JM | j_{BC}; J' M' \rangle = 0 \quad \forall J \neq J' \text{ or } M \neq M'. \quad (4)$$

Bases of this type are known as Jordan bases of subspaces [14]. Since a state of the first basis (labeled by j_{AB}) has overlap with only one state of the second basis (labeled by j_{BC}), the problem is reduced to a discrimination instance between pairs of pure states. Then the total error probability is simply the sum of the contributions of each pair.

In the unambiguous approach, the minimum probability of an inconclusive result for a pair of states $|\phi_1\rangle, |\phi_2\rangle$

with equal priors is simply $P^?(|\phi_1\rangle, |\phi_2\rangle) = |\langle\phi_1|\phi_2\rangle|$ [1], hence

$$P^? = \frac{1}{d_{AB}d_C} \sum_{JM} |\langle j_{AB}; JM | j_{BC}; JM \rangle|. \quad (5)$$

These overlaps can be computed in terms of the Wigner 6- j symbols [24]:

$$\langle j_{AB}; JM | j_{BC}; JM \rangle = (-1)^{j_A+j_B+j_C+J} \sqrt{(2j_{AB}+1)(2j_{BC}+1)} \begin{Bmatrix} j_A & j_B & j_{AB} \\ j_C & J & j_{BC} \end{Bmatrix}, \quad (6)$$

and they are independent of M [24], therefore in what follows we omit writing the quantum number M , and we perform the sum over M in (5) trivially by adding the multiplicative factor $2J+1$. Substituting the value of the 6 j symbols for $j_A = j_C = n/2$, $j_B = m/2$, $j_{AB} = j_{BC} = (n+m)/2$, and setting $J = m/2 + k$ we obtain

$$\langle j_{AB}; J | j_{BC}; J \rangle = \binom{n}{k} \binom{n+m}{n-k}^{-1}, \quad (7)$$

with $k = 0, 1, \dots, n$ (observe that J takes values from $J = n + m/2$ of the totally symmetric space down to $J = m/2$).

Plugging the overlaps Eq. (7) into Eq. (5), we obtain

$$\begin{aligned} P^? &= \sum_{k=0}^n \frac{m+2k+1}{(n+m+1)(n+1)} \frac{(m+k)!n!}{(m+n)!k!} \\ &= 1 - \frac{nm}{(n+1)(m+2)}, \end{aligned} \quad (8)$$

where notice that the dimension of the subspace of total angular momentum J is $m+2k+1$ and in the second equality we have used the binomial sums

$$\begin{aligned} \sum_{k=0}^n \binom{m+k}{m} &= \binom{n+m+1}{m+1}, \\ \sum_{k=0}^n k \binom{m+k}{m} &= \binom{n+m+1}{m+1} \frac{n(m+1)}{m+2}. \end{aligned} \quad (9)$$

In the minimum error approach no inconclusive results are allowed, but the machine is permitted to give wrong answers with some probability that one tries to minimize. This minimum error probability can be computed along the same lines as in the previous case. Recall that the error probability P^{ME} for two pure states $|\phi_1\rangle, |\phi_2\rangle$ and equal a priori probabilities is [2]

$$P^{\text{ME}}(|\phi_1\rangle, |\phi_2\rangle) = \frac{1}{2} \left(1 - \sqrt{1 - |\langle\phi_1|\phi_2\rangle|^2} \right). \quad (10)$$

The total error probability is just the sum of the contribution of each pair of states with the same quantum

numbers JM , $\{|j_{AB}; JM\rangle, |j_{BC}; JM\rangle\}$,

$$\begin{aligned} P^{\text{ME}} &= \frac{1}{2} \left(1 - \sum_{k=0}^n \frac{m+2k+1}{(n+1)(n+m+1)} \right. \\ &\quad \left. \times \sqrt{1 - \left(\frac{(m+k)!n!}{(m+n)!k!} \right)^2} \right). \end{aligned} \quad (11)$$

It is instructive to obtain the well-known results when the ports are loaded with just one copy of each state [9], i.e., $n = m = 1$.

The inconclusive probability in the unambiguous approach reads

$$\begin{aligned} P^? &= \frac{1}{6} \sum_{J=1/2}^{3/2} (2J+1) |\langle j_{AB} = 1; J | j_{BC} = 1; J \rangle| \\ &= \frac{5}{6}, \end{aligned} \quad (12)$$

i.e., five out of six times the machine gives an inconclusive result and only 1/6 of the times identifies the state without error. Notice that the overlaps for $J = 3/2$ are one. This must be so since $J = 3/2$ corresponds to the totally symmetric subspace, which is independent of the way the spins coupled. That is, this subspace is identical for σ_1 and σ_2 . This is the main source of error as it contributes $4/6 = 4/6 \times 1$ out of the total $5/6$ error probability. The remaining $1/6 = 2/6 \times 1/2$ is the contribution of the $J = 1/2$ subspace, where the $2/6$ is the probability of having an outcome on this subspace and $1/2$ is the overlap between the states [cf. Eq. (7)].

The minimum error probability in the one copy case reads

$$\begin{aligned} P^{\text{ME}} &= \frac{1}{2} \left(1 - \frac{1}{6} \sum_{J=1/2}^{3/2} (2J+1) \right. \\ &\quad \left. \times \sqrt{1 - |\langle j_{AB} = 1; J | j_{BC} = 1; J \rangle|^2} \right), \end{aligned} \quad (13)$$

which using Eq. (7) or directly Eq. (11) gives

$$P^{\text{ME}} = \frac{1}{2} \left(1 - \frac{1}{2\sqrt{3}} \right) \simeq 0.356. \quad (14)$$

That is, approximately 1/3 of the times the outcome of the machine will be incorrect.

The error probability in both, minimum error and unambiguous approaches, will of course decrease when using more copies of the states at the ports of the discrimination machine. Equations (8) and (11) give the unambiguous and minimum error probability for arbitrary values of n and m . They enable us to study the behaviour of the machine for a large number of copies in the program and the data ports, which is what we next discuss.

III. ASYMPTOTIC LIMITS FOR PURE STATES

Let us start by considering the case of an asymptotically large number of copies at the program ports ($n \rightarrow \infty$) while keeping finite the number of copies m at the data port. For the unambiguous discrimination one obtains from Eq. (8)

$$\lim_{n \rightarrow \infty} P^? = \frac{2}{m+2}. \quad (15)$$

We wish to show that in this limit the programmable machine has a performance that is equivalent to a protocol consisting in first estimating the states and then doing a discrimination of *known* states. The average of the inconclusive probability of this protocol over all input states should coincide with Eq. (15). Recall that for known $|\psi_1\rangle$ and $|\psi_2\rangle$ states, when a number m of copies of the unknown state is given, this probability reads

$$P^?(\psi_1, \psi_2) = |\langle \psi_1 | \psi_2 \rangle|^m. \quad (16)$$

One can do an explicit calculation of the average $\langle P^?(\psi_1, \psi_2) \rangle = 1/2 \int_0^\pi \sin \theta \cos^m \theta / 2$, but it is amusing to obtain it in a very simple way from the Schur lemma

$$\begin{aligned} \int d\psi_2 \left(|\langle \psi_1 | \psi_2 \rangle|^2 \right)^{\frac{m}{2}} &= \langle \psi_1 |^{\otimes \frac{m}{2}} \left(\int d\psi_2 |\psi_2\rangle^{\otimes \frac{m}{2}} \right) |\psi_1\rangle^{\otimes \frac{m}{2}} \\ &= \frac{1}{d_{m/2}} = \frac{1}{m/2 + 1}, \end{aligned} \quad (17)$$

where $d_{m/2}$ is the dimension of the symmetric space of $m/2$ qubits (notice that *stricto sensu* this procedure is only valid for m even). Plugging this average into (16) one immediately recovers Eq. (15).

Now we turn our attention to the minimum error probability. Taking $n \rightarrow \infty$ and using the Stirling approximation $z! \approx z^z e^{-z} \sqrt{2\pi z}$ in Eq. (11), one obtains

$$\begin{aligned} \lim_{n \rightarrow \infty} P^{\text{ME}} &= \frac{1}{2} \left[1 - 2 \int_0^1 dx x \sqrt{1 - x^{2m}} \right] \\ &= \frac{1}{2} \left[1 - \frac{\sqrt{\pi}}{2} \frac{\Gamma(1 + 1/m)}{\Gamma(3/2 + 1/m)} \right], \end{aligned} \quad (18)$$

where we have defined $x = k/n$ and used the Euler-McLaurin summation formula at leading order $\sum_{k=0}^n f(k) \simeq n \int_0^1 dx f(nx)$.

This result could be easily anticipated from the minimum error probability with classical knowledge of the pure states. Recall that the minimum error probability given m identical copies is

$$P^{\text{ME}}(\psi_1, \psi_2) = \frac{1 - \sqrt{1 - |\langle \psi_1 | \psi_2 \rangle|^{2m}}}{2}, \quad (19)$$

and we just have to compute the average for all pairs of the above expression. Using $|\langle \psi_1 | \psi_2 \rangle|^2 = \cos^2 \theta / 2$, where θ is the relative angle between the Bloch vectors of the two states, one has

$$\langle P^{\text{ME}}(\psi_1, \psi_2) \rangle = \frac{1}{2} \left[1 - \frac{1}{2} \int_0^\pi d\theta \sin \theta \sqrt{1 - \cos^{2m}(\theta/2)} \right], \quad (20)$$

and performing the change of variables $x = \sin \theta / 2$ this equation is cast exactly in the form of (18).

What can not be anticipated is the next order $O(1/n)$, which gives the very relevant information on how fast the protocol reaches the asymptotic value (18). A lengthy, but rather straightforward, calculation yields the remarkable result that this term has a coefficient which coincides with the value of the integral $\int_0^1 dx x \sqrt{1 - x^{2m}}$. At this order we therefore can write

$$P^{\text{ME}} = \frac{1}{2} - \frac{\sqrt{\pi}}{4} \frac{\Gamma(1 + 1/m)}{\Gamma(3/2 + 1/m)} \left(1 - \frac{1}{n} \right). \quad (21)$$

We now analyze the complementary case, that is, when the number of copies at the data port is infinitely large, $m \rightarrow \infty$, while the number n of copies at the program ports is kept finite. In this limit we have perfect knowledge of the data state, $|\psi\rangle$, but we do not know to which program port it should be associated. Observe that this situation is very much the same as state comparison [19].

In this scenario the inconclusive probability in the unambiguous approach reads from (8)

$$\lim_{m \rightarrow \infty} P^? = \frac{1}{n+1}. \quad (22)$$

Let us see that this agrees with the average performance of standard state comparison. If the data state is the same as the program state in the upper or lower port, the effective states to be discriminated are

$$\begin{aligned} \sigma_1 &= \frac{1}{d_n} [\psi^{\otimes n}] \otimes \mathbb{1}_n \\ \sigma_2 &= \frac{1}{d_n} \mathbb{1}_n \otimes [\psi^{\otimes n}], \end{aligned} \quad (23)$$

respectively, where $d_n = n+1$ is the dimension of the symmetric space of n -qubits and $\mathbb{1}_n$ is the projector onto this subspace. The minimal inconclusive probability for these two states can be obtained with a POVM whose elements are $\{E_1 = [\psi^{\otimes n}] \otimes [\psi^{\otimes n}]^\perp, E_2 = [\psi^{\otimes n}]^\perp \otimes [\psi^{\otimes n}]^\perp\}$

$[\psi^{\otimes n}], E_? = \mathbb{1} \otimes \mathbb{1} - E_1 - E_2\}$, where $[\psi^{\otimes n}]^\perp = \mathbb{1}_n - [\psi^{\otimes n}]$, that is, with a POVM that checks in each register whether the state is $|\psi\rangle$ or not. Then the inconclusive probability reads

$$P^?(\psi) = \frac{1}{2} (\text{tr} E_? \sigma_1 + \text{tr} E_? \sigma_2) = \frac{1}{n+1} \quad (24)$$

independently of the state $|\psi\rangle$, where we have used $\text{tr}[\psi^{\otimes n}]^\perp = \text{tr}(\mathbb{1}_n - [\psi^{\otimes n}]) = n$.

The minimum error probability in this limit can be tackled in a similar fashion. The asymptotic expression of Eq. (11), though not as direct as in the unambiguous case, is rather straightforward to obtain. Notice that the dominant factor in the term containing factorials inside the square root is $m^{-2(n-k)}$. So, we can effectively replace the square root term by one, for all $k < n$. Taking into account that for $k = n$ the square root vanishes, we have

$$\lim_{m \rightarrow \infty} P^{\text{ME}} = \frac{1}{2} \left(1 - \frac{n}{n+1} \right) = \frac{1}{2(n+1)}. \quad (25)$$

The minimum error probability of a strategy that first estimates perfectly the input state and then tries to associate the correct label to it is given by Helstrom formula for σ_1 and σ_2 [2]

$$P^{\text{ME}} = \frac{1}{2} (1 - \frac{1}{2} \|\sigma_1 - \sigma_2\|), \quad (26)$$

where $\|A\| = \text{tr} \sqrt{A^\dagger A}$ is the trace-norm of operator A . Substituting the expression of the states (23) we obtain

$$\begin{aligned} P^{\text{ME}} &= \frac{1}{2} \left(1 - \frac{1}{2(n+1)} \|[\psi^{\otimes n}] \otimes [\psi^{\otimes n}]^\perp \right. \\ &\quad \left. - [\psi^{\otimes n}]^\perp \otimes [\psi^{\otimes n}]\| \right) \\ &= \frac{1}{2} \left(1 - \frac{2}{2(n+1)} \|[\psi^{\otimes n}] \otimes [\psi^{\otimes n}]^\perp\| \right) \\ &= \frac{1}{2} \left(1 - \frac{n}{n+1} \right) = \frac{1}{2(n+1)}, \end{aligned} \quad (27)$$

where in the first equality we have subtracted the common term $[\psi^{\otimes n}] \otimes [\psi^{\otimes n}]$ from both states, in the second we have used the orthogonality of the operators and in the last equality we use $\text{tr}[\psi^{\otimes n}]^\perp = n$. As expected, the result is again independent of $|\psi\rangle$.

To end this section we compute the asymptotic error probabilities for the symmetric case, that is, when all the ports are loaded with the same $m = n$ (and large) number of copies.

In the unambiguous approach when $n = m \rightarrow \infty$ the first nonvanishing order of (8) reads

$$P^? = \frac{3}{n} + \dots \quad (28)$$

To compute the minimum error probability, it is convenient to write Eq. (11) for $n = m$ as

$$P^{\text{ME}} = \frac{1}{2} \sum_{k=0}^n p_k \left(1 - \sqrt{1 - c_k^2} \right), \quad (29)$$

where

$$p_k = \frac{n+1+2k}{(2n+1)(n+1)} \quad (30)$$

and

$$c_k = \frac{\binom{n+k}{n}}{\binom{2n}{n}}. \quad (31)$$

We first observe that c_k is a monotonically increasing function and hence it takes its maximum value at $k = n$. Second, we note that around this point

$$\begin{aligned} \binom{n+k}{n} &\simeq 2^{(n+k)H(\frac{n}{n+k})} \\ &\simeq 2^{(n+k)H(1/2)} = 2^{n+k}, \end{aligned} \quad (32)$$

where $H(x) = -x \ln x - (1-x) \ln(1-x)$ is the Shannon entropy of a binary random variable and we have used that $k \approx n$ and $H(1/2) = 1$. Similarly, one has

$$\binom{2n}{n} \simeq 2^{2nH(1/2)} = 2^{2n} \quad (33)$$

and hence $c_k \simeq 2^{-(n-k)}$. With this, the probability of error in this limit reads

$$P^{\text{ME}} = \frac{1}{2} \sum_{k=0}^{\infty} p_k \left(1 - \sqrt{1 - \left(\frac{1}{4} \right)^{n-k}} \right). \quad (34)$$

Finally, we perform the change of variables $k \rightarrow n-k$ and use that in Eq. (30) $p_{n-k} \simeq 3/(2n)$ for $k \simeq 0$ to obtain

$$P^{\text{ME}} = \frac{3}{4n} \zeta(1/4) \approx \frac{0.882}{n}, \quad (35)$$

where we have defined the function

$$\zeta(x) = \sum_{k=0}^{\infty} \left(1 - \sqrt{1 - x^k} \right), \quad (36)$$

which converges very quickly to its exact value (the first four terms already give a value that differ in less than 10^{-3} from the exact value).

IV. MIXED STATES

We now move to the case when the program and data ports are loaded with mixed states. This situation arises when, e.g., there are imperfections in the preparation or noise in the transmission of the states. It is reasonable to suppose that these imperfections have the same effect on all states, i.e. to consider that the states have all the same purity r . The input states are then tensor products of

$$\rho_i = \frac{\mathbb{1} + r \vec{n}_i \vec{\sigma}}{2}, \quad (37)$$

where \vec{n}_i is a unitary vector and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the usual Pauli matrices. In what follows we assume that only the purity is known, i.e. one knows the characteristics of the noise affecting the states, but nothing else. This means that the averages will be performed over the isotropic Haar measure of the \mathbb{S}^2 sphere, like for pure states. At the end of this section we also analyze the performance of a fully universal discrimination machine, that is, when not even the purity is considered to be known.

Notice that mixed states can only be unambiguously discriminated if they have different supports [21], which is not the case when the ports are loaded with copies of the states (37) as they are full-rank matrices. Therefore, only the minimum error discrimination approach will be analyzed here. It is worth stressing that the computation of the optimal error probability in the multi-copy case is very non-trivial, even for known qubit mixed states. Only recently feasible methods for computing the minimum error probability for a rather large number of copies have been developed and the asymptotic expression of this probability has been obtained [4, 6]. The main difficulty can be traced back to the computation of the trace-norm [see Eq.(26)] of large matrices. The dimension of the matrices grows exponentially with the total number of copies entering the machine, and for a relative small number of them the problem becomes unmanageable. However, as it will be clear, it is possible to exploit the permutation symmetry of the input states to write them in block-diagonal form [22, 23] crucially reducing the complexity of the problem.

The two effective states we have to discriminate are

$$\begin{aligned}\sigma_1 &= \int dn_1 dn_2 \rho_{1A}^{\otimes n} \otimes \rho_{1B}^{\otimes m} \otimes \rho_{2C}^{\otimes n} \\ \sigma_2 &= \int dn_1 dn_2 \rho_{1A}^{\otimes n} \otimes \rho_{2B}^{\otimes m} \otimes \rho_{2C}^{\otimes n},\end{aligned}\quad (38)$$

where $dn_i = d\Omega_i/(4\pi)$ is the invariant measure on the two-sphere. Any state having permutation invariance, as e.g. $\rho^{\otimes n}$, can be written in a block diagonal form using the irreducible representations of the symmetric group S_n . Each block is specified by the total angular momentum j and a label α that distinguishes the different equivalent representations for a given j

$$\rho^{\otimes n} = \bigoplus_{j,\alpha} \rho^{j\alpha}. \quad (39)$$

The angular momentum takes values $j = n/2, n/2 - 1, \dots, 1/2(0)$ for odd (even) n and the number of equivalent representations for each j is [23]

$$\nu_j^n = \binom{n}{n/2-j} \frac{2j+1}{n/2+j+1}, \quad (40)$$

That is $\alpha = 1, \dots, \nu_j^n$. For each block we have [23]

$$\begin{aligned}\text{tr} \rho^{j\alpha} &= \left(\frac{1-r^2}{4}\right)^{n/2-j} \sum_{k=-j}^j \left(\frac{1-r}{2}\right)^{j-k} \left(\frac{1+r}{2}\right)^{j+k} \\ &\equiv (2j+1)C_j^n,\end{aligned}\quad (41)$$

which, of course, is the same for all equivalent irreducible representations, i.e., independent on the label α . We sketch here the origin of the factors appearing in (41) (full details can be found in [23]). The first factor comes from the contribution from the $n/2 - j$ singlets present in a representation j made up of n spin-1/2 states. The summation term is the trace of the projection of the remaining states in the symmetric subspace with total angular momentum j , where we can use the rotational invariance of the trace to write each state in diagonal form $\begin{pmatrix} \frac{1+r}{2} & 0 \\ 0 & \frac{1-r}{2} \end{pmatrix}$. This term simply reads

$$\begin{aligned}t_j &= \sum_{k=-j}^j \left(\frac{1-r}{2}\right)^{j-k} \left(\frac{1+r}{2}\right)^{j+k} \\ &= \frac{1}{r} \left[\left(\frac{1+r}{2}\right)^{2j+1} - \left(\frac{1-r}{2}\right)^{2j+1} \right]\end{aligned}\quad (42)$$

and hence

$$C_j^n = \frac{1}{2j+1} \left(\frac{1-r^2}{4}\right)^{n/2-j} t_j. \quad (43)$$

Very much in the same way as it happened in previous sections, the only difference between the diagonal basis of σ_1 and σ_2 is the ordering of the angular momenta couplings. In σ_1 we first couple subspaces A and B and obtain

$$\rho_{AB} = \int dn_1 \rho_{1A}^{\otimes n} \otimes \rho_{1B}^{\otimes m} = \sum_{\xi_{AB}} C_{j_{AB}}^{n+m} \mathbb{1}_{\xi_{AB}}, \quad (44)$$

where

$$\mathbb{1}_{\xi_{AB}} = \sum_{M_{AB}} |\xi_{AB} M_{AB}\rangle \langle \xi_{AB} M_{AB}| \quad (45)$$

is the projector onto the subspace with quantum numbers $\xi_{AB} = \{j_A, \alpha_A, j_B, \alpha_B, j_{AB}\}$ and $C_{j_{AB}}^{n+m}$ is defined in Eq. (41). Notice that $C_{j_{AB}}^{n+m}$ depends only on the purity of the state and on the total angular momentum j_{AB} . Notice also that the tensor product of a mixed state has projections in all subspaces and the blocks are not uniquely determined by the value of j_{AB} , i.e., one has to keep track of the labels j_A and j_B as well. Of course, subspaces with different quantum numbers ξ_{AB} are orthogonal, i.e., $\text{tr}[\mathbb{1}_{\xi} \mathbb{1}_{\xi'}] = \delta_{\xi\xi'} \text{tr} \mathbb{1}_{\xi}$. When coupling the third system one plainly adds the quantum numbers $\xi_C = \{j_C, \alpha_C\}$.

In the notation we have developed so far, the diagonal bases of σ_1 and σ_2 are written as $\mathcal{B}_1 = \{|\xi_{AB} \xi_C; JM\rangle\}$

and $\mathcal{B}_2 = \{|\xi_A \xi_{BC}; JM\rangle\}$, respectively. Obviously, each set contains 2^{2n+m} orthonormal states and Eq. (38) reads

$$\begin{aligned}\sigma_1 &= \sum_{\xi_{AB}\xi_C} \sum_{JM} C_{j_{AB}}^{n+m} C_{j_C}^n [\xi_{AB}\xi_C; JM] \\ \sigma_2 &= \sum_{\xi_A\xi_{BC}} \sum_{JM} C_{j_A}^n C_{j_{BC}}^{n+m} [\xi_A\xi_{BC}; JM].\end{aligned}\quad (46)$$

We just have to compute the minimum error from the Helstrom formula (26) for these two states. It is convenient to define the trace-norm term

$$T = \|\sigma_1 - \sigma_2\|, \quad (47)$$

so that

$$P^{\text{ME}} = \frac{1}{2} \left(1 - \frac{1}{2} T \right). \quad (48)$$

To compute T we need to know the unitary matrix Λ that transforms \mathcal{B}_2 into \mathcal{B}_1 or vice versa. The elements of this unitary are given by the overlaps between the elements of both basis $\langle \xi_{AB}\xi_C; JM | \xi'_A \xi'_{BC}; J'M' \rangle$. We observe that these overlaps are non-vanishing only if $j_X = j'_X$, $\alpha_X = \alpha'_X$ ($X = A, B, C$) and $J = J', M = M'$. Furthermore, as mentioned above, their value does not depend on M or α_X , thus, sums over these quantum numbers simply amount to introduce the corresponding multiplicative factors. Therefore, it is useful to introduce a label containing the quantum numbers that determine the orthogonal blocks in \mathcal{B}_1 and \mathcal{B}_2 that may have non vanishing overlaps, $\xi = \{j_A, j_B, j_C, J\}$ and the corresponding multiplicative factor

$$\gamma_\xi = \nu_{j_A}^n \nu_{j_B}^m \nu_{j_C}^n (2J+1), \quad (49)$$

where ν_j^n is given in Eq. (40). Eq. (47) then reads

$$T = \sum_{\xi} \gamma_{\xi} T^{\xi} = \sum_{\xi} \gamma_{\xi} \|\sigma_1^{(\xi)} - \Lambda^{(\xi)} \sigma_2^{(\xi)} \Lambda^{(\xi)T}\|, \quad (50)$$

where the explicit expressions of the matrix elements are

$$\begin{aligned}(\sigma_1^{(\xi)})_{j_{AB}j'_{AB}} &= \delta_{j_{AB}j'_{AB}} C_{j_{AB}}^{n+m} C_{j_C}^n \\ (\sigma_2^{(\xi)})_{j_{BC}j'_{BC}} &= \delta_{j_{BC}j'_{BC}} C_{j_A}^n C_{j_{BC}}^{n+m}\end{aligned}\quad (51)$$

and

$$\Lambda_{j_{AB}, j_{BC}}^{(\xi)} = \langle \xi, j_{AB} | \xi, j_{BC} \rangle. \quad (52)$$

Recall that the overlap (52) is independent of the quantum number labelling the equivalent representations (recall also that it is independent of M) and therefore is given by Eq. (6).

The computation of the minimum error probability reduces to a sum of trace-norms of small size Helstrom matrices that have dimensions of the allowed values of j_{AB} and j_{BC} for given $\xi = \{j_A, j_B, j_C, J\}$. Hence

$$P^{\text{ME}} = \frac{1}{2} \left(1 - \frac{1}{2} \sum_{\xi} \gamma_{\xi} T^{\xi} \right) \quad (53)$$

and this computation can be done very efficiently.

We would like to show the analytical results for the simplest case of having just one state at each port, i.e. when $n = m = 1$. In this situation we have fixed values $j_A = j_B = j_C = 1/2$, the total angular momentum can be $J = 3/2, 1/2$ and $j_{AB} = 1, 0$ (and similarly for j_{BC}). Here there is no degeneracy, the number of equivalent representations defined in Eq. (40) is one, and therefore the multiplicative factor (49) simply reads $\gamma_{\xi} = 2J+1$. The only relevant quantum number in this case is $\xi = J$, as all the other are fixed, and we do not need to write them explicitly. The minimum error probability is then

$$P^{\text{ME}} = \frac{1}{2} \left(1 - \frac{1}{2} \sum_{J=1/2}^{3/2} (2J+1) \|\sigma_1^{(J)} - \Lambda^{(J)} \sigma_2^{(J)} \Lambda^{(J)T}\| \right) \quad (54)$$

The term of the sum corresponding to $J = 3/2$ vanishes since it corresponds to the projection of $\sigma_{1,2}$ onto the completely symmetric subspace, which is identical for both states. Indeed, in this subspace $\sigma_1^{(3/2)} = \sigma_2^{(3/2)} = C_1^2 C_{1/2}^1 = (3+r^2)/24$, where we have used Eq. (43), and from Eq. (52) we obtain $\Lambda^{(3/2)} = 1$. In the subspace $J = 1/2$ we have

$$\begin{aligned}\sigma_1^{(1/2)} = \sigma_2^{(1/2)} &= \begin{pmatrix} C_1^2 C_{1/2}^1 & 0 \\ 0 & C_0^2 C_{1/2}^1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{24} (3+r^2) & 0 \\ 0 & \frac{1}{8} (1-r^2) \end{pmatrix}\end{aligned}\quad (55)$$

and

$$\Lambda^{(1/2)} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}. \quad (56)$$

Plugging these expressions into (54) we obtain the minimum error probability of one copy state

$$P^{\text{ME}} = \frac{1}{2} \left(1 - \frac{r^2}{2\sqrt{3}} \right). \quad (57)$$

As expected, when $r \rightarrow 1$ we recover the pure state value (14).

Numerical results of the minimum error probability as a function of the purity of the input states for the symmetric case $n = m$ are depicted in Fig. 1. One sees that for low values of n ($n \lesssim 3$) the dependence on the purity is not very marked, the curves are concave almost in the whole range of the purity. For larger n however there is an interval of purities where the behavior changes quite significantly. For, e.g., $n = 29$ the inflection point occurs at $r \approx 0.3$. At very large values of n one expects a step-like shape with an inflection point approaching $r = 0$ because the probability of error remains very small for $r \neq 0$ and is strictly $1/2$ at $r = 0$. The shape of the curves is explained by the existence of two distinct regimes. For high purities the probability of error is well fitted by a linear function in the inverse of the number of

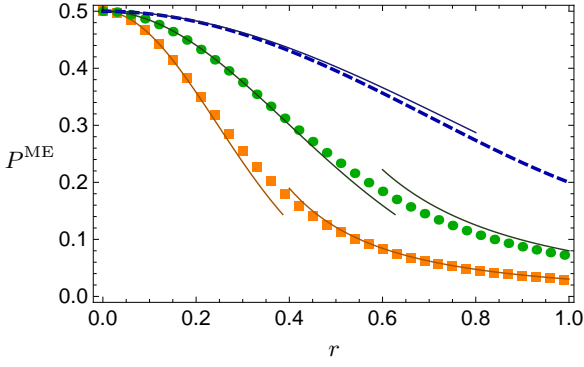


FIG. 1. (color online) Error probability P^{ME} for $n = m = 3$ (blue dashed line), 11 (green circles) and 29 (yellow squares) versus purity. The fit $P^{\text{ME}} \simeq 0.882/(nr^2)$ in the regime of high purities for $n = 11$ and $n = 29$ and the Gaussian approximation $P^{\text{ME}} \simeq 1/2 \exp[-nr^2/(2\sqrt{3})]$ in the regime of low purities for all cases is represented (solid lines).

copies. We get $P^{\text{ME}} \simeq 0.88/(nr^2)$ where the value 0.88 coincides with the analytical value computed for pure states Eq. (35). Of course, this approximation cannot be valid for low purities. In this range of low purity the minimum error probability is very well approximated by the Gaussian function $P^{\text{ME}} \simeq 1/2 \exp[-nr^2/(2\sqrt{3})]$, where we have taken the argument of the exponential from the exponentiation of the exact $1 \times 1 \times 1$ case (57). This approximation works for purities in the interval of the width of the Gaussian, i.e., up to $\sim 1/\sqrt{n}$. Therefore, as n increases the asymptotic approximation $P^{\text{ME}} \propto 1/(nr^2)$ extends its validity to almost the whole range of purities, and the expected jump discontinuity develops in $r = 0$ as $n \rightarrow \infty$. Similar information is depicted in Fig. 2, where the error probability is plotted as function of the number of copies n for different purities. We have superimposed the asymptotic result, which is seen to yield a very good approximation to the exact error probability already for $n \gtrsim 20$.

V. ASYMPTOTIC $n \times 1 \times n$

As in previous sections, it is interesting to study the performance of the machine in the asymptotic regimes. A particularly important instance where it is possible to obtain closed expressions is the case when the number of copies at the program ports is asymptotically large and there is one state at the data port. We show how to compute the leading order and sketch the generalizations needed to obtain the subleading term.

Observe first that j_{AB} can only take the values $j_{AB} = j_A \pm 1/2$ and similarly for j_{BC} . Therefore $\sigma_{1,2}^{(\xi)}$ are 2×2 matrices (except in the extremal case of $J = j_A + j_C + 1/2$ which is one-dimensional). It is useful to write

$$\sigma(j) = C_{j_A}^n C_{j_C}^n \begin{pmatrix} R_+(j) & 0 \\ 0 & R_-(j) \end{pmatrix} \quad (58)$$

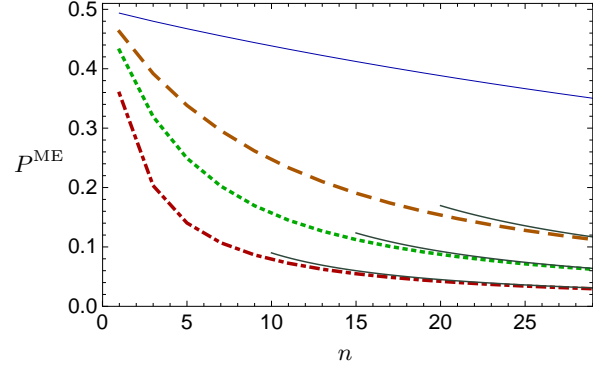


FIG. 2. (color online) Error probability P^{ME} for $r = 0.2$ (blue thin solid line), $r = 0.5$ (brown dashed line), $r = 0.7$ (green dotted line) and $r = 1$ (red dot-dashed line) versus n ($n = m$ is assumed). Numerical points have been joined for an easier visualization. The approximation $0.882/(nr^2)$ is also represented (thin solid lines).

with

$$R_{\pm}(j) = \frac{C_{j \pm 1/2}^{n+1}}{C_j^n}. \quad (59)$$

With this definition one simply has [see Eq.(51)]:

$$\sigma_1^{(\xi)} = \sigma(j_A) \quad \text{and} \quad \sigma_2^{(\xi)} = \sigma(j_C). \quad (60)$$

We further notice that for large n

$$\begin{aligned} \nu_j^n C_j^n &\approx \frac{1}{n/2 + j + 1} \frac{1+r}{2r} \\ &\times \sqrt{\frac{2}{n\pi(1-r^2)}} \exp \left[-n \frac{(2j/n - r)^2}{2(1-r^2)} \right]. \end{aligned} \quad (61)$$

Defining $y = 2j/n$ and using the Euler-Maclaurin summation formula, we have for a generic function $f(j)$

$$\sum_j \nu_j^n C_j^n f(j) \approx \frac{1+r}{2r} \int_{-\infty}^{\infty} \frac{dy G_n(y)}{n/2 + ny/2 + 1} f\left(\frac{ny}{2}\right), \quad (62)$$

where we have extended limits of integration from $(0,1)$ to $(-\infty, \infty)$, which is legitimate for large n , and defined

$$G_n(y) = \sqrt{\frac{n}{2\pi(1-r^2)}} \exp \left[-n \frac{(y-r)^2}{2(1-r^2)} \right], \quad (63)$$

a Gaussian distribution centered at $y = r$ with variance $\sigma^2 = (1-r^2)/n$. Notice that at leading order, $n \rightarrow \infty$, $G_{\infty} \approx \delta(y-r)$ and hence

$$\sum_j \nu_j^n C_j^n f(j) \approx \frac{1}{nr} f\left(\frac{nr}{2}\right). \quad (64)$$

Notice also that at this order

$$R_{\pm}(j) \approx R_{\pm}\left(\frac{nr}{2}\right) = \frac{1 \pm r}{2}. \quad (65)$$

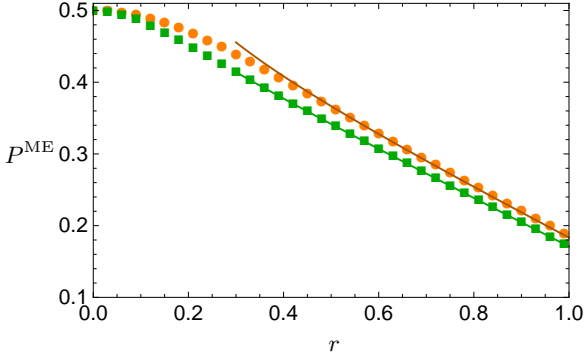


FIG. 3. (color online) Error probability P^{ME} for $n = 20$ (yellow circles) and $n = 79$ (green squares) versus purity. The asymptotic behaviour given by Eq. (70) is represented for both cases.

There only remains to compute the unitary matrix Eq. (52). Observe that the total angular momentum takes values $J = |j_A - j_C| + 1/2 + k$ with $k = 0, 1, \dots, 2 \min\{j_A, j_C\}$. The leading order is rather easy to write (the subleading term, although straightforward, is far more involved, and we will not show it here). At this order we have $J = 1/2 + k$ and $k = 0, 1, \dots, nr$ and the matrix elements computed from (6) yield

$$\Lambda^{(\xi)} = \frac{1}{nr} \begin{pmatrix} k & \sqrt{(nr)^2 - k^2} \\ \sqrt{(nr)^2 - k^2} & -k \end{pmatrix}. \quad (66)$$

Plugging Eqs.(58-66) into Eq. (50) one gets

$$T \simeq \sum_{k=0}^{nr} 2k \times \frac{2}{n^3 r^2} \sqrt{(nr)^2 - k^2} \quad (67)$$

where the sum over j_A and j_C has been trivially performed by substituting their central value $nr/2$ in the summand and the only remaining multiplicative of γ_ξ [cf. Eq. (49)] is $2J + 1 \simeq 2k$. Finally, defining $x \equiv k/nr$ and using the Euler Mac-Laurin approximation as in Eq. (18) we obtain

$$T \simeq 4r \int_0^1 dx x \sqrt{1 - x^2} = \frac{4r}{3}, \quad (68)$$

and hence

$$P^{\text{ME}} \simeq \frac{1}{2} - \frac{r}{3}, \quad (69)$$

which obviously coincides with the pure state result Eq. (18) for $m = 1$ and $r \rightarrow 1$.

As for the computation of the next-to-leading order, the integrals approximating the sums over j_A and j_C have to incorporate the fluctuations around the central value, i.e., one defines $j_A = \frac{n}{2}(r + \eta_A)$ and $j_C = \frac{n}{2}(r + \eta_C)$, where the variables η_X have effective dimension $n^{-1/2}$. Then one can expand the matrix elements of $\sigma_{1,2}$, Λ , and the terms of ν_j^n present in Eq. (62) taking into account the effective dimensionality of all the terms [notice that $k \rightarrow n(r + \eta)x$, where the integration range of x is

$(0, 1)$]. One then performs the sum in k by means of the Euler-Maclaurin summation formula as before. Finally one computes the integration in $j_{A/B}$ taking into account that range of the variables $\eta_{A/B}$ can be taken to be $(-\infty, \infty)$. After a somewhat lengthy calculation we obtain

$$P^{\text{ME}} \simeq \frac{1}{2} - \frac{r}{3} + \frac{1}{3nr}. \quad (70)$$

Notice that the limit $r = 0$ is singular and not surprisingly the expansion breaks down for purities of order $1/n$. As it should, the error probability (70) increases monotonically with the purity.

In Fig. 3 we plot the error probability as a function of the purity for $n = 20$ and $n = 79$. One sees that the asymptotic expression (70) approximates very well the minimum error probability even for small number of copies. For larger n , e.g., for $n = 79$ the approximation works extremely well down to values below $r = 0.3$.

We finish this section by showing that the leading term (69) coincides with the average error of a device that first estimates the mixed states at the program ports and afterwards does the usual minimum error discrimination of the data state. From the Helstrom formula (26) particularized for mixed qubit states one has

$$P^{\text{ME}} = \left\langle \frac{1}{2} \left(1 - \frac{1}{2} |\vec{r}_1 - \vec{r}_2| \right) \right\rangle \quad (71)$$

where the average is taken over all possible orientations of the Bloch vectors \vec{r}_1 and \vec{r}_2 . For equal purity states it simply reads

$$P^{\text{ME}} = \frac{1}{2} \left(1 - \frac{r}{2} \int_0^\pi d\theta \sin \theta \sin \theta/2 \right) = \frac{1}{2} - \frac{r}{3}. \quad (72)$$

VI. FULLY UNIVERSAL DISCRIMINATION MACHINE

Let us finally address the fully universal discrimination machine, i.e., a machine that distinguishes states from which nothing is assumed to be known, not even its purity. For this type of machine, we need to specify a prior distribution for the purity. While the isotropy of the angular variables yields a unique uniform distribution for the angular variables, the Haar measure on the 2-sphere used in previous sections, the corresponding expression for a fully unbiased distribution of the purity $w(r)$ is not uniquely determined. This is a longstanding issue, and several priors haven been suggested depending on the assumptions made [7, 26]. Here we will not stick to a particular distribution, rather we will show results for three reasonable distributions. The actual values of the the probability of error may depend on the chosen prior, but the overall performance is seen to be very similar.

The most straightforward, but perhaps not very well grounded, choice is that of the distribution of a hard-sphere $w(r) \propto r^2$, i.e., a normalized integration measure

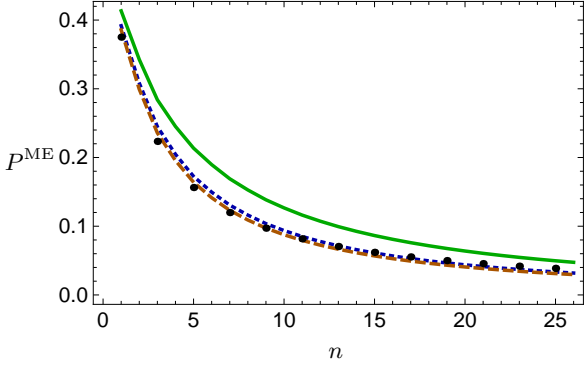


FIG. 4. (color online) Error probability P^{ME} for hard-sphere (green solid line), Bures (blue dotted line) and Chernoff (red dashed line) priors versus n ($n = m$ is assumed). The points correspond to the error probability for a fixed $r = 0.9$; its proximity to the Chernoff curve exposes the fact that this prior gives larger weights to states of high purity.

given by

$$d\rho^{\text{HS}} = 3r^2 dr \frac{d\Omega}{4\pi}. \quad (73)$$

The Bures distribution is far better motivated. It corresponds to the volume element induced by the fidelity distance [25]. It is monotonically decreasing under coarse graining [26] and it has been argued that it corresponds to maximal randomness of the signal states [27]. In this case one has $w(r) \propto r^2/\sqrt{1-r^2}$. Notice that this distribution assigns larger weights to pure states, as their distinguishability in terms of the fidelity is larger than that of mixed states. The integration measure reads

$$d\rho^{\text{Bu}} = \frac{4}{\pi} \frac{r^2}{\sqrt{1-r^2}} dr \frac{d\Omega}{4\pi}. \quad (74)$$

Lastly, we also consider the recently proposed Chernoff distribution [4]. It is the prior induced by the Chernoff distance, which has a clear operational meaning in terms of the distinguishability between states. By construction it is monotonically decreasing under coarse graining. This measure assigns even larger weights to states of high purity and lower to the very mixed ones. This assignment is again based on the distinguishability properties, but in terms of the asymptotic behavior of the error probability. The measure can be written as [4]

$$d\rho^{\text{Ch}} = \frac{1}{\pi - 2} \frac{(\sqrt{1+r} - \sqrt{1-r})^2}{\sqrt{1-r^2}} dr \frac{d\Omega}{4\pi}. \quad (75)$$

The effective states we have to discriminate are

$$\Sigma_k = \int d\rho_1 d\rho_2 \rho_{1A}^{\otimes n} \otimes \rho_{kB}^{\otimes m} \otimes \rho_{2C}^{\otimes n} \quad (k = 1, 2), \quad (76)$$

where $d\rho_k$ takes the expressions of the measures (73-75). Note that the block structure of the states is the same as before, as it only depends on the permutation invariance of the input states, which remains untouched. Further,

we can use rotational invariance in the same fashion as in Eqs. (44) and (46). Therefore, here it is only required to compute the average of the coefficients C_j^n in Eq. (41) according to priors (73-75). To calculate the minimum error probability of this fully universal machine one simply uses Eq. (53) for the states (46) with the averaged coefficients $\langle C_j^n \rangle$ computed in Appendix B.

In Fig. 4 we present the minimum error probability of the fully universal machine for the three priors discussed for an equal number of program and data states up to $n = m = 26$. As anticipated, the smaller average error corresponds to the Chernoff distance, because states with higher purity are assigned a larger weight, and these are easier to discriminate. The probability of error, as somehow expected, is inversely proportional to the number of copies, and attains very similar values than for the discrimination of states with fixed known purity of the order of $r \sim 0.9$.

VII. CONCLUSIONS

We have studied the problem of programmable discrimination of two unknown general qubit states when multiple copies of the states are provided. For pure states we have obtained the optimal unambiguous discrimination and minimum-error probabilities, Eqs. (8) and (11), respectively. Some results along these lines can be found in [17], however no closed expressions were given there. Knowing the error in the asymptotic regimes is a very relevant information as it allows to assess and compare the performance of devices in a way which is independent on the number of copies. We have obtained analytical expressions for the leading and sub-leading terms in several cases of interest. As could be anticipated, when the number of copies at the program ports is asymptotically large at leading order we recover the average of the usual discrimination problem of known states Eqs. (17) and (20). When the data port is loaded with an asymptotically large number of copies, we recover the state comparison averaged error Eqs. (24) and (27). These cases correspond to measure and discriminate protocols, where the measurement unveils the classical information about the states.

We have addressed for the first time the programmable discrimination of copies of mixed states. We have obtained the minimum-error probability when the ports are loaded with copies of qubits of known purity, Eq. (53). We have assumed that all states have the same purity. This would, e.g., correspond to a scenario where all the initially pure data and program states are subject to the same depolarizing noise before entering the machine. Closed analytical results for a small number of copies can be obtained and efficiently computable expressions for a fairly large number of copies are given. The asymptotic analytical results show very good agreement with the numerics. The latter show a characteristic $1/N$ dependence with the number N of available copies—in contrast

to the usual exponential decay found in standard (non-universal) state discrimination—and provide a very good approximation already for relatively low number of copies when the states have high purity. For very mixed states the error probability has a drastically different behavior. Logically, in both cases the error probability monotonically decreases with increasing purity r , but in the low purity regime the dependence is much less pronounced. The range of purities exhibiting this behavior shrinks as the number of copies increases, and the characteristic $1/N$ behavior of the asymptotic regime extends its validity over almost the whole range of purities.

Finally we have studied the fully universal discrimination machine, a device that takes in states of which nothing is known, i.e., not even its purity. We compute the minimum error probability for three reasonable prior distributions of the purity: hard-sphere, Bures and Chernoff (see Fig. 4). The latter is seen to give the lowest error probability. This comes as no surprise, since the Chernoff distribution assigns larger weights to pure states (because they are better distinguished). Our results also indicate that the fully universal discrimination states yields an error probability comparable to the discrimination of states of known purity, being that remarkably large ($r \sim 0.9$).

ACKNOWLEDGMENTS

We thank J. Bergou for their contributions in the earlier stages of this work. We acknowledge financial support from: the Spanish MICINN, through the Ramón y Cajal program (JC), FPI grant BES-2009-028117 (GS), contract FIS2008-01236, and project QOIT (CONSOLIDER2006-00019); from the Generalitat de Catalunya CIRIT, contract 2009SGR-0985. We thank M. Hayashi for pointing out a discrepancy with the wrong Eq. (A2) in the published version of our paper.

Appendix A: Arbitrary copies of pure states

In this appendix we present the unambiguous discrimination and minimum error probabilities when the number of copies n_A, n_B, n_C loaded at the machine ports is completely arbitrary. Note that in this case the global states σ_1 and σ_2 [cf. Eq. (2)] may have different dimensions, for $d_1 = (n_A + n_B + 1)(n_C + 1)$ is in general not equal to $d_2 = (n_A + 1)(n_B + n_C + 1)$. One can easily convince oneself that the support of the state with smallest dimension is always contained in the support of the other, and hence the problem can be solved in very much the same way as in the main text simply taking into account that the error probabilities now only contain contributions from the intersection of the supports. Without loss of generality we can assume from now on

that $n_A \geq n_C$. As discussed in the main text, the error probabilities are computed by adding the pairwise contributions of the state bases in the common support, the main difference being that σ_1 and σ_2 do not have equal coefficients in front of the projectors and hence the prior probabilities of each pair of states are different. Also, the overlaps in Eq. (6) will have a slightly more complicated expression. Here we have $j_A = n_A/2$, $j_B = n_B/2$, $j_C = n_C/2$, $j_{AB} = (n_A + n_B)/2$ and $j_{BC} = (n_B + n_C)/2$. The minimum J available for σ_1 is $j_B + j_A - j_C \equiv J_{min}^1$, and $|j_B + j_C - j_A| \equiv J_{min}^2$ for σ_2 . The maximum angular momentum $j_A + j_B + j_C \equiv J_{max}$ is reachable for both states. For equal prior probabilities for σ_1 and σ_2 , we can write

$$\frac{1}{2}\sigma_1 = \sum_{J=J_{min}^1}^{J_{max}} \sum_{M=-J}^J p_J \pi_J^1[j_{AB}; JM], \quad (A1)$$

$$\frac{1}{2}\sigma_2 = \sum_{J=J_{min}^2}^{J_{max}} \sum_{M=-J}^J p_J \pi_J^2[j_{BC}; JM], \quad (A2)$$

where $p_J = \frac{1}{2} \left(\frac{1}{d_1} + \frac{1}{d_2} \right)$, $\pi_J^1 = \frac{1}{2p_J d_1}$, $\pi_J^2 = \frac{1}{2p_J d_2}$ for $J_{min}^1 \leq J \leq J_{max}$, whereas $p_J = \frac{1}{2d_2}$, $\pi_J^1 = 0$, $\pi_J^2 = 1$ for $J_{min}^2 \leq J < J_{min}^1$. We can view p_J as the probability of projecting on subspace labeled by a given J and M , and π_J^1, π_J^2 as the prior probabilities that, on that subspace, the unknown state be $[j_{AB}; JM]$ or $[j_{BC}; JM]$. If the condition

$$\frac{c_J^2}{1 + c_J^2} \leq \pi_J^{AB} \leq \frac{1}{1 + c_J^2}, \quad (A3)$$

where $c_J = |\langle j_{AB}; JM | j_{BC}; JM \rangle|$ is given by Eq. (6), holds, then the probability of obtaining an inconclusive answer when we finally discriminate between $[j_{AB}; JM]$ and $[j_{BC}; JM]$ is [28] $P_J^? = 2\sqrt{\pi_J^1 \pi_J^2} c_J$. If Eq. (A3) is satisfied for $\hat{J} = J_{max} - 1$, then it will be satisfied all over this range of J , since c_J is a monotonically increasing function of J . The overlap c_J has the very simple form

$$c_J^2 = \frac{n_A n_C}{(n_A + n_B)(n_B + n_C)}. \quad (A4)$$

Thus, Eq. (A3) is equivalent to

$$\begin{aligned} \frac{n_A n_C}{(n_A + n_B)(n_B + n_C)} &\leq \frac{(n_A + n_B + 1)(n_C + 1)}{(n_B + n_C + 1)(n_A + 1)} \\ &\leq \frac{(n_A + n_B)(n_B + n_C)}{n_A n_C}, \end{aligned} \quad (A5)$$

which is clearly true. Eq. (A3) does not hold if $J = J_{max}$, for which we have $P_{J_{max}}^? = 1$. Notice that since no error is made for $J_{min}^2 \leq J < J_{min}^1$, for $\pi_J^1 = 0$, the total inconclusive probability reads $P^? = \sum_{J=J_{min}^1}^{J_{max}} p_J (2J + 1) P_J^?$, which has the explicit expression

$$P^? = \frac{1}{2} \left(\frac{1}{\sqrt{d_1}} - \frac{1}{\sqrt{d_2}} \right)^2 d_{ABC} + \frac{1}{\sqrt{d_1 d_2}} \sum_{k=0}^{n_C} (n_A + n_B - n_C + 2k + 1) \sqrt{\frac{\binom{n_A + n_B - n_C + k}{n_B} \binom{n_B + k}{n_B}}{\binom{n_A + n_B}{n_B} \binom{n_C + n_B}{n_B}}}, \quad (\text{A6})$$

where $d_{ABC} = n_A + n_B + n_C + 1$. Note that when $n_A = n_C$ the term proportional to d_{ABC} disappears and the square root term simplifies, so we recover the closed form given in the main text [cf. Eq. (8)].

The minimum error probability can be computed entirely along the same lines. For a pair of states we have [28] $P_J^{\text{ME}} = \frac{1}{2} \left(1 - \sqrt{1 - 4\pi_j^1 \pi_j^2 c_j^2} \right)$, and the total error probability reads

$$P^{\text{ME}} = \frac{1}{4} \left\{ 1 + \frac{d_1}{d_2} - \frac{d_1 + d_2}{d_1 d_2} \sum_{k=0}^{n_C} (n_A + n_B - n_C + 2k + 1) \sqrt{1 - 4 \frac{d_1 d_2}{(d_1 + d_2)^2} \frac{\binom{n_A + n_B - n_C + k}{n_B} \binom{n_B + k}{n_B}}{\binom{n_A + n_B}{n_B} \binom{n_C + n_B}{n_B}}} \right\}. \quad (\text{A7})$$

This expression coincides with Eq. (31) of [29].

For the hard sphere prior we have

$$\begin{aligned} \langle C_j^n \rangle_{\text{HS}} &= 3 \int C_j^n r^2 dr \\ &= 6 \frac{\Gamma(n/2 + j + 2) \Gamma(n/2 - j + 1)}{\Gamma(n + 4)}. \end{aligned} \quad (\text{B2})$$

Appendix B: Averaged C_j^n coefficients

Here we compute the average of the coefficients [see Eq. (41)]

The Bures distribution yields

$$\begin{aligned} \langle C_j^n \rangle_{\text{Bu}} &= \frac{4}{\pi} \int C_j^n \frac{r^2}{\sqrt{1 - r^2}} dr \\ &= \frac{4}{\pi} \frac{\Gamma(n/2 + j + 3/2) \Gamma(n/2 - j + 1/2)}{\Gamma(n + 3)}. \end{aligned} \quad (\text{B3})$$

The averages for the Chernoff prior are a bit more involved, but still can be given in a closed form as

$$\begin{aligned} C_j^n &= \frac{1}{2j+1} \left(\frac{1-r^2}{4} \right)^{n/2-j} \\ &\times \sum_{k=-j}^j \left(\frac{1-r}{2} \right)^{j-k} \left(\frac{1+r}{2} \right)^{j+k} \end{aligned} \quad (\text{B1})$$

$$\begin{aligned} \langle C_j^n \rangle_{\text{Ch}} &= \frac{1}{\pi - 2} \int C_j^n \frac{(\sqrt{1+r} - \sqrt{1-r})^2}{\sqrt{1-r^2}} dr \\ &= \frac{2}{(\pi - 2)(2j+1)} \sum_{m=-j}^j \left[B_{1/2} \left(\frac{n+1-2m}{2}, \frac{n+1+2m}{2} \right) \right. \\ &\quad \left. - 2B_{1/2} \left(\frac{n-2m+2}{2}, \frac{n+2m+2}{2} \right) \right] \end{aligned} \quad (\text{B4})$$

where $B_x(a, b) = \int_0^x t^{a-1} (1-t)^{b-1} dt$ is the incomplete beta function [30].

for the hard sphere, Bures and Chernoff priors, Eqs. (73-75), considered in the fully universal discrimination machine.

- [1] I. D. Ivanovic, Phys. Lett. A **123**, 257 (1987); D. Dieks, Phys. Lett. A **123**, 257 (1987); A. Peres, Phys. Lett. A **128**, 19 (1988).
- [2] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
- [3] A. Hayashi, T. Hashimoto, and M. Horibe, Phys. Rev. A **78**, 012333 (2008); H. Sugimoto, T. Hashimoto, M. Horibe, and A. Hayashi, Phys. Rev. A **80**, 052322 (2009).
- [4] K. M. R. Audenaert *et al.*, Phys. Rev. Lett. **98**,

- 160501 (2007); J. Calsamiglia *et al.*, Phys. Rev. A **77**, 032311 (2008).
- [5] M. Nussbaum and A. Szkola, Ann. Statist. **37** 1040, (2009).
- [6] J. Calsamiglia, J. I. de Vicente, R. Muñoz-Tapia, and E. Bagan, Phys. Rev. Lett. **105**, 080504 (2010).
- [7] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States* (Cambridge University Press, Cambridge, UK, 2006).
- [8] M. Dusek and V. Buzek, Phys. Rev. A **66**, 022112 (2002).

- [9] J. A. Bergou and M. Hillery, Phys. Rev. Lett. **94**, 160501 (2005).
- [10] E. Bagan, S. Iblisdir, and R. Muñoz-Tapia, Phys. Rev. A **73**, 022341 (2006).
- [11] M. Guta and W. Kotłowski, arXiv: 1004.2468. For machine-learning related ideas in the context of quantum information science see also: Aïmeur E., Brassard G. and Gambs, Proc. 19th Canadian Conference on Artificial Intelligence (Canadian AI'06), pp. 433-444. Springer-Verlag; A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, P. Perinotti, Phys. Rev. A **81**, 032324 (2010); S. Hentschel and B. C. Sanders, Phys. Rev. Lett. **104**, 063603 (2010).
- [12] M. Sedlak, M. Ziman, O. Pribyla, V. Buzek, and M. Hillery, Phys. Rev. A **76**, 022326 (2007); M. Sedlak, M. Ziman, V. Buzek, and M. Hillery, Phys. Rev. A **79**, 062305 (2009).
- [13] L. Bartuskova, A. Cernoch, J. Soubusta, and M. Dusek, Phys. Rev. A **77**, 034306 (2008).
- [14] J. A. Bergou, E. Feldman, and M. Hillery, Phys. Rev. A **73**, 032107 (2006).
- [15] A. Hayashi, M. Horibe, and T. Hashimoto, Phys. Rev. A **72**, 052306 (2005); A. Hayashi, M. Horibe, and T. Hashimoto, Phys. Rev. A **73**, 012328 (2006).
- [16] J. Bergou, V. Buzek, E. Feldman, U. Herzog, and M. Hillery, Phys. Rev. A **73**, 062334 (2006).
- [17] B. He and J. A. Bergou, Phys. Rev. A **75**, 032316 (2007).
- [18] U. Herzog and J. Bergou, Phys. Rev. A **78**, 032320 (2008), Erratum: Phys. Rev. A **78**, 069902 (2008).
- [19] S. M. Barnett, A. Chefles, and I. Jex, Phys. Lett. A **307**, 189 (2003).
- [20] M. Sedlak *et al.*, Phys. Rev. A **77**, 042304 (2008).
- [21] T. Rudolph, R. W. Spekkens, and T. S. Turner, Phys. Rev. A **68**, 022308 (2003); P. Raynal, N. Ltkenhaus, and S. J. van Enk, Phys. Rev. A **68**, 022308 (2003); U. Herzog and J. Bergou, Phys. Rev. A **71**, 050301(R) (2005).
- [22] J. I. Cirac, A. K. Ekert, and C. Macchiavello, Phys. Rev. Lett. **82**, 4344 (1999).
- [23] E. Bagan, M. A. Ballester, R. D. Gill, A. Monras, and R. Muñoz-Tapia, Phys. Rev. A **73**, 032301 (2006).
- [24] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1960)
- [25] K. Zyczkowsky and H. S. Sommers, Phys. Rev. A **71**, 032313 (2005).
- [26] D. Petz and C. Sudar, J. Math. Phys. **37**, 2662 (1996).
- [27] M. J. W. Hall, Phys. Lett. A **242**, 123 (1998).
- [28] J.A. Bergou *et al.*, Lect. Notes Phys. **649**, 417465 (2004).
- [29] D. Akimoto and M. Hayashi, arXiv:1102.2555, unpublished
- [30] M. Abramowitz and I. A. Stegun, eds., *Handbook of Mathematical Functions* (Dover Publications, New York, 1972).